AIMS VIDEO COURSES

SUPPORTING BOOKLET

PROBABILITY & STATISTICS

WITH

PROF DAVID SPIEGELHALTER

AIMS

SOUTH AFRICA

AIMS
SOUTH AFRICA

African Institute for Mathematical Sciences

6 MELROSE ROAD | MUIZENBERG | CAPE TOWN 7945 | SOUTH AFRICA
TEL: +27 (0)21 787 9320 | FAX: +27 (0)21 787 9321
EMAIL: info@aims.ac.za | WEB: www.aims.ac.za

AIMS Online Courses

The mission of the AIMS academic programme is to provide an excellent, advanced education in the mathematical sciences to talented African students in order to develop independent thinkers, researchers and problem solvers who will contribute to Africa's scientific development.

Teaching at AIMS is based on the principle of learning and understanding, rather than simply listening and writing, during classes, and on creating an atmosphere of increasing our knowledge through class discussions, through small group discussions, by formulating conjectures and assessing the evidence for them, and sometimes going down wrong paths and learning from the mistakes that led us there. The essential features of the classes at AIMS are that, in contrast to formal lecture courses, they are highly interactive, where the students engage with the lecturer throughout the class time, are encouraged to learn together in a journey of questioning and discovery, and where lecturers respond to the needs of the class rather than to a pre-determined syllabus. AIMS teaching philosophy is to promote critical and creative thinking, to experience the excitement of learning from true understanding, and to avoid rote learning directed only towards assessment.

Leading international and local experts offer the courses at AIMS, which are three weeks long (each module consisting of 30 hrs) and collectively form the coursework for a structured masters degree which also includes a research component. The advertised content is a guide, and the lecturers are encouraged, and indeed expected, to adapt daily to meet the current needs of the students.

Over the past ten years AIMS has achieved international recognition for this innovative and flexible approach. It has been the starting point for the remarkable success of our students and alumni and we all benefit from the support of many who have "witnessed the AIMS-magic and keep coming back for more."

This year we have decided to film selected courses and to make them available to a larger audience as an online facility. African universities may choose to use these courses to supplement and enhance their own postgraduate programmes. We believe this would be best achieved through engagement with AIMS. One way for this to happen, would be for AIMS to suggest or nominate a specialist tutor to spend time at the university, guiding students who follow the online programme. Where possible expert lecturers who have taught at AIMS may visit the university to give a short introduction to the course. We would welcome this interaction as well as the contribution our online courses will make to the growth of the mathematical sciences ecosystem in Africa.

Barry Green
Director & Professor of Mathematics
African Institute for Mathematical Sciences
January 2013

PROBABILITY & STATISTICS

2012

PROF DAVID SPIEGELHALTER

DAY 2

AIMS

SOUTH AFRICA

Outline solutions and hints - 1

1. *I flip a coin 4 times. What is the probability of getting an odd number of 'Heads'?*

   Draw a tree ending in 16 possible outcomes, such as H, H, H, H, H, H, H, T etc . Each split has probability 0.5. Multiplying down each branch gives joint probability of each outcomes as 1/16. Adding up over branches with an odd number of Heads give 8/16 = 0.5.

2. *A drawer contains 3 white socks and 3 black socks. If I pick 2 socks at random (without replacement), what is the chance I get a* non-matching *pair?*

   Draw sequentially: Probability of sock 1 being white is 0.5, Prob(sock 2 is white | sock 1 white)= 2/5 = 0.4, so probability that both are white = 0.5 x 0.4 = 0.2. Complete tree to find probability non-matching pair is 0.6.

3. *If I put the first sock back in the drawer before taking the second, how does this change the probability of a non-matching pair?*

   Events are now independent: chance of non-matching pair is now 0.5 (all sequences equally likely)

4. *If I throw 2 dice 24 times, what is the chance of getting at least one 'double-6'?*

   The answer is NOT 24/36 (what if we carried out the experiment 100 times?): this is the *expected* number of double-6's

   Aagin consider 'six / not-six' split at each throw. we need

   = 1 - (probability of NOT getting a double-6) [*complement rule*]

   = 1 - (probability of not getting a double-6 on 1st throw AND not on the 2nd throw AND ... AND not on 24th throw) [*multiplication rule for independent events*]

   = 1 - $\left(\frac{35}{36}\right)^{24}$ = 1 - 0.51 = 0.49. So the Chevalier had spotted that this bet was slightly against him.

5. *Write a program in SciPy to simulate throwing a dice 4 times, repeat this experiment 1,000 times and see in what proportion a '6' appeared*

   Version using Numpy

```
from __future__ import division
import numpy as np
import random


N = 1000;         #Number of time you want to do the experiment
Ne=4                #Number of times we need to throw the die in each experiment
num_six = 0    #Initialise results
for j in range(N):     #Carry out experiment N times
         results = np.zeros(Ne)     #Initialise intermediate result as an empty array
         for i in range(Ne):             # Loop to throw 4 times
```

```
                throw = random.randint(1,6) # Throw the die once and saving the result
                results[i]=throw
        if any(results==6): # Checking to see if any of the 4 throws are = 6
                num_six = num_six+1

#Calculate the probability of getting a 6 = total number of successful events/ total number
Prob = num_six/N
True_prob = 1 - (5/6)**4 # From class calculation
print "The numerical probability is: ",Prob, "The true probability is: ", True_prob
```

Version using SciPy

```
from __future__ import division
from scipy import random
N = 1000;  #Number of times you want to do the experiment
num_six = 0   #Initialise results
for j in range(N):   #Carry out experiment N times
         throw = random.randint(1,7,4) # throwing 4 times at once
         if any(throw==6): # Checking to see if there are any 6's
                  num_six = num_six+1
```

6. *Write a program in SciPy to simulate throwing two dice 24 times, repeat this experiment 1,000 times and see in what proportion of sequences a 'double-6' appeared. Try 1,000,000 times. Compare with the exact answer (see above)*

```
from __future__ import division
from scipy import random
N = 1000;  #Number of time you want to do the experiment
Nt = 24   #Number of throws
success=0   #Initialise results
for j in range(N):   #Carry out experiment N times
         num_six_pairs = 0   # Initialising the results
         throw = random.randint(1,7,(2,Nt))   # Throwing the 2 dice, 24 times at once
         for i in range(Nt): # Checking how many pairs of 6's
                  if all(throw[:,i]==6):
                           num_six_pairs+=1
         if num_six_pairs>0: #Since we need only one success, noting that
                  success+=1
#Calculate the probability of getting a 6
Prob = success/N
#True probability from theory
True_prob = 1 - (35/36)**24
print "The numerical probability is: ",Prob, "The true probability is: ", True_prob
```

7. *A college in a very foreign country is composed of 70% men and 30% women. It is known that 40% of the men, and 60% of the women, smoke cigarettes. If I see someone smoking*

a cigarette, what is the probability that it is a woman? Under what circumstances would this be a realistic assessment of a probability?!

Could think of what we would expect of 100 random students: 70 men, of whom 28 smoke and 42 don't, and 30 women, of whom 18 smoke and 12 don't. So there would be 46 smokers, of whom 18 are women, so the probability is $18/46 = 0.39$. This would only be reasonable if I can assume the person I see is equally likely to be any individual, whether male or female (e.g. not at a football match)

8. ** A drawer contains white socks and black socks. if I pick 2 socks at random, the probability of picking 2 white socks is 0.5. What is the smallest number of socks for which this is true? Find another possible solution.*

Assume w white socks, b black socks, sampling without replacement and so we have the identity

$$\frac{w}{w+b} \times \frac{w-1}{w+b-1} = \frac{1}{2}$$

Can rearrange to form a Diaphantine equation (integer solution). One solution is fairly obvious. There is another, which is not so obvious. I have no idea whether there are more!

2: Equally-likely events: more formal

1. Probability as enumerable ratio
2. Permutations
3. Combinations
4. Alternative ways of calculation
5. Coincidences

Probability as enumerable ratio

Suppose an experiment $\mathcal{E}$ has n possible outcomes, judged equally likely.

For a set A with cardinality r, then

$$P(A) = \frac{\text{Ways that satisfy } A}{\text{Total number of possibilities}} = \frac{r}{n}.$$

So working out probabilities becomes a problem of enumeration

Dice: $A =$ odd numbers $1, 3, 5$, $P(A) = 3/6 = 0.5$.

Ordered sets

Number of ordered r-tuples where each chosen from $1, 2, ., n$ is n^r

Cardano throwing 2 dice: number of *distinct ordered outcomes* is $6 \times 6 = 36$

Equivalent to *sampling with replacement*

- A box has n balls labelled $1, ..., n$
- Take a ball, note its number, *put it back*
- After r draws, number of possible ordered sequences is n^r

Permutations: sampling without replacement

- Same set up but ball *is not put back*
- After r draws, number of possible ordered sequences is $n(n-1)..(n-r+1) = \frac{n!}{(n-r)!} = {}^nP_r$
- If $n = r$, then ${}^nP_r = n!$
- $n!$ = total number of permutations: the number of ways of putting n balls in n boxes

Birthday coincidences

Assume people are equally likely to have any of $n = 365$ birthdays (ignore February 29th, family planning etc!)

If there are r people in a room, what is the chance they all have different birthdays? i.e. event $D_r = b_1 \neq b_2 \neq ... \neq b_r$

$$
\begin{aligned}
P(D_r) &= \frac{\text{No. ways of choosing birthdays that are all different}}{\text{Total no. ways choosing birthdays}} = \frac{{}^{n}P_r}{n^r} \\
&= \frac{365!}{r!}/365^r \\
&= \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)..\left(1 - \frac{r-1}{365}\right)
\end{aligned}
$$

For $n = 23$, this is 0.48

So only need 23 people in a room to have greater than 50% chance of two having the same birthday

Alternative ways of calculation

There are often at least 2 ways of calculating probabilities: and it is not obvious which is the best way!

Using sequential probability trees is generally more intuitive

$$P(D_r) = P(b_2 \neq b_1 \text{ AND } b_3 \neq (b_1, b_2)....) = \frac{364}{365} \times \frac{363}{365} ... \times \frac{366 - r}{365}$$

Combinations

How many distinct *unordered* subsets of size r are there from n objects?

There are ${}^nP_r = n!/(n-r)!$ ordered subsets (sampling without replacement)

Each subset can be arranged in $r!$ ways, and so each unordered subset is represented $r!$ times

And so there are $\frac{n!}{r!(n-r)!}$ denoted ${}^nC_r$ or $\left(\begin{array}{c} n \\ r \end{array} \right)$

If a lottery picks 6 numbers from 1 to 49, how many possible distinct tickets are there? What is the chance of winning the jackpot?

Can do this using combinatoric formula, or using a sequential probability tree.

3: Discrete random variables (more formal)

1. Sample spaces etc
2. Random variable as a function
3. Expectation and variance
4. Bernoulli, Binomial, Poisson, Geometric distributions

Sample spaces etc

- For an experiment $\mathcal{E}$, define a sample space Ω of possible outcomes ω
- We can consider that 'nature' chooses an observed ω^{obs} according to some probability law P
- We say that an 'event' A has occurred if and only if $\omega^{\text{obs}} \in A$
- $P(A)$ is the probability that A will occur for a future experiment

σ-fields

A nonempty collection of subsets $\mathcal{A}$ of a set Ω is called a σ-field provided that

1. If A is in $\mathcal{A}$, then A^c is also in $\mathcal{A}$
2. If A_n is in $\mathcal{A}$, $n = 1, 2, 3, ..,$ then $\cup_{n=1}^{\infty} A_n$ and $\cap_{n=1}^{\infty} A_n$ are both in $\mathcal{A}$

Restricts the subsets of $\mathcal{A}$ over which we shall define a probability distribution.

Probability spaces

A probability measure P on a σ-field of subsets of $\mathcal{A}$ of a set Ω is a real-valued function having domain $\mathcal{A}$ for which

1. $P(\Omega) = 1$
2. $P(A) \geq 0$ for all A in $\mathcal{A}$
3. If A_n is in $\mathcal{A}$, $n = 1, 2, 3, ..,$ are mutually disjoint sets in $\mathcal{A}$, then $P\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$

A probability space is denoted by $(\Omega, \mathcal{A}, \mathcal{P})$.

Measure theory concerns more general functions on σ-fields - basis for integration on continuous domains.

Example: flip coin twice:

$\Omega = HH, HT, TH, TT$, $\mathcal{A}$ is set of all subsets of Ω, P is defined from $P(\omega_i) = 0.25$ for each $\omega_i \in \Omega$

Conditional probability

Definition: Let A and B be two events with $P(A) > 0$. Then the conditional probability of B given A, written $P(B|A)$, is defined to be

$$P(B|A) = \frac{P(B \cap A)}{P(A)}.$$

If $P(A) = 0$, $P(B|A)$ is undefined (pointless to condition on an impossible event)

From these basic rules, all else follows!

For example,

- Complement rule: can prove $P(A) + P(A^c) = P(A \cup A^c) = P(\Omega) = 1$
- Addition rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (create disjoint sets)

Random variable as a function

A discrete valued random variable X on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ is a function X from Ω to the real numbers R

The real-valued function f defined on R by $f_X(x) = P(X = x)$ is called the *discrete density function* (probability mass function)

$F_X(x) = P(X \leq x)$ is called the *discrete distribution function* : the event $X \leq x$ is interpreted as $\omega : X(\omega) \leq x$

Example: flip coin twice, let X be the total number of heads H

Then

$$\begin{aligned} f_X(0) &= P(X = 0) = P(TT) = 0.25 \\ f_X(1) &= P(X = 1) = P(HT \cup TH) = 0.50 \\ f_X(2) &= P(X = 2) = P(HH) = 0.25 \end{aligned}$$

Expectation and variance

For a discrete random variable X with a finite set of possible values $x_1, .., x_r$, we define

- Expectation (mean): $\mathbb{E}_X[X] = \sum_{i=1}^{r} x_i f(x_i)$
- Variance: $\mathbb{V}_X[X] = \sum_{i=1}^{r} (x_i - \mathbb{E}_X[X])^2 f(x_i)$
- Standard deviation $= \sqrt{\text{Variance}}$

Generally easier to compute (dropping suffix)

$$\mathbb{V}[X] = \sum_{i=1}^{r} x_i^2 f(x_i) - (\mathbb{E}[X])^2 = \mathbb{E}_X[X^2] - \mathbb{E}[X]^2$$

since $\sum_{i=1}^{r} (x_i - \mathbb{E}[X])^2 f(x_i) = \sum_{i=1}^{r} [x_i^2 - 2\mathbb{E}[X]\, x_i + \mathbb{E}[X]^2] f(x_i) = \mathbb{E}[X^2] - 2\sum_{i=1}^{r} \mathbb{E}[X]\, x_i f(x_i) + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

When there is a parameter, say p, then conditioning can be explicitly represented as $f_X(x|p)$ etc

Bernoulli distribution

The distribution of an 'indicator' 0/1 variable :

Bernoulli trial with probability p of 'success'

- Density: $f_X(0|p) = 1 - p; f_X(1|p) = p$
- Expectation (mean): $\mathbb{E}_X[X|p] = (1-p) \times 0 + p \times 1 = p$
- Variance:
  $\mathbb{V}_X[Xp] = \mathbb{E}_X[X^2] - \mathbb{E}_X[X]^2 = (1-p) \times 0^2 + p \times 1^2 - p^2 = p(1-p)$
- Standard deviation $= \sqrt{p(1-p)}$

Binomial distribution

The distribution of the sum of n Bernoulli trials

How many successes out of n trials, each with probability p of success?

Probability of a particular sequence of x successes and $n - x$ failures is $p^x(1-p)^{n-x}$

But there are $\frac{n!}{x!(n-x)!} = \left(\begin{array}{c} n \\ x \end{array} \right)$ such sequences

- Denoted: Binomial(n, p)
- Density: $f_X(x|p) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x(1-p)^{n-x}$; $x = 0, 1, 2, 3, .., n$
- Expectation (mean): $\mathbb{E}_X[X|p] = np$
- Variance: $\mathbb{V}_X[X|p] = np(1-p)$
- Standard deviation $= \sqrt{np(1-p)}$

[will see how to get this mean and variance later]

Questions - 2

1. Check you are happy with the code for the previous questions on throwing dice.

2. A group of 50 students want to elect a committee. In how many ways can a president, vice-president, secretary and treasurer be chosen? In how many ways could a committee of 4 people, who will share all the tasks, be chosen?

3. 20 people in a room choose an integer at random from $1, 2, .., 100$. What is the chance at least 2 people choose the same number?

4. You throw a dice 6 times and count the number X of 6's that appear. What is the name for the distribution of X? What is the probability density function for X?

5. Using the random.binomial function to throw the 6 dice simultaneously, adapt the program for throwing dice to simulate 1000 replications from the distribution of X. Compare to that expected from the theoretical distribution from the previous distribution